Detecting the Trustworthiness of Novel Partners in Economic Exchange

David DeSteno[1], Cynthia Breazeal[2], Robert H. Frank[3], David Pizarro[3], Jolie Baumann[1], Leah Dickens[1], Jin Joo Lee[2]

[1]Northeastern University
[2]Massachusetts Institute of Technology Media Lab
[3]Cornell University

In press, *Psychological Science*

Author Note

David DeSteno, Jolie Baumann, and Leah Dickens, Department of Psychology, Northeastern University. Cynthia Breazeal and Jin Joo Lee, Media Lab, Massachusetts Institute of Technology; Robert Frank and David Pizarro, Cornell University.

We acknowledge the support of NSF research grants BCS 0827084/88/94.

Correspondence concerning this manuscript should be addressed to David DeSteno at the following address:

David DeSteno
Dept. of Psychology
Northeastern University
Boston, MA 02115
Voice: 617.373.7884
Fax: 617.373.8714
Email: d.desteno@gmail.com

Draft Date: March 17, 2012

Abstract

Deciding to trust a person one knows nothing about often entails high risk. The ability to infer a potential partner's trustworthiness would consequently be highly advantageous. To date, however, little evidence supports the ability of humans to accurately assess the cooperative intentions of novel partners through the use of nonverbal signals. In two studies using human-human and human-robot interactions, we show that accuracy in judging the trustworthiness of novel partners is heightened through exposure to nonverbal cues as well as identify a specific set of cues as predictive of economic behavior. Employing the precision offered by robotics technology to model and control human-like movements, we demonstrate not only that experimental manipulation of the identified cues directly impacts perceptions of trustworthiness and subsequent exchange behavior but also the readiness of the human mind to utilize such cues to ascribe social intentions to technological entities.

Word Count: 143

People must often decide whether to trust new potential partners in the absence of reliable information about past behavior (Axelrod, 1984; Delton, Krasnow, Cosmides, & Tooby, 2011; Frank, 1988). Although forming an entirely new exchange relationship can be extremely advantageous, decisions to trust also entail risks of substantial loss if one's partner acts in an untrustworthy manner. Consequently, any capacity that enhances accuracy in detecting the trustworthiness of others would offer a significant competitive advantage.

In the absence of reliable information about an individual's reputation, one source of possible information about another's likely actions may stem from nonverbal cues. Supporting this view, ample evidence exists that humans regularly use specific cues, often without conscious awareness, to infer the motivations of others with some level of accuracy (Ambady & Weisbuch, 2010; Knapp & Hall, 2010). To date, however, the nature of the cues that might predict trustworthy or untrustworthy behavior, or even whether such cues exist, remains unclear. Yet, for cooperation to occur to the degree it does in humans, it appears theoretically necessary that information related to reliable, albeit imperfect, signals of trustworthy intent or to the likelihood of subsequent encounters with potential partners be available (Delton et al, 2011; Frank, 1988).[1] Indeed, without access to trust-relevant signals in situations where reputational information is absent (e.g., interactions with strangers) and where the likelihood of repeated interactions is low, it is likely that the advantages of acting opportunistically would reduce cooperation substantially.

Given the theoretical import and adaptive advantages of an ability to assess trustworthiness, the search for trust-relevant signals has long occupied the attention of scholars from many fields (e.g., psychology, behavioral economics, evolutionary biology). Researchers have looked in vain for a single dynamic "golden cue" that predicts whether a person can actually be trusted (Ambady & Weisbuch, 2010; Knapp & Hall, 2010). In a similar vein, researchers have also looked for and identified certain markers in static faces that impact judgments of trustworthiness (Todorov, Baron & Oosterhof, 2008); however, there is little reliable evidence linking such markers to the actual behaviors of others (Todorov, 2008).[2]

We believe that past difficulties in identifying relevant signals may stem from attempts to look at cues individually or in isolation from social interaction. We suspect that if reliable trust-related signals are to be found, they will likely be ones that emerge dynamically and are processed intuitively within the context of interpersonal situations. Given that each member of a dyad is attempting to assess the intentions of an unfamiliar other, signals will likely be subtle and unfold over time as individuals evaluate their potential partner. As Frank (1988) notes, interpreting cooperative signals must necessarily entail some level of uncertainty, as a perfectly readable signal would have long since relegated opportunistic individuals to extinction.

Given increasing work suggesting that the interpretation of nonverbal cues is highly context-dependent (Ambady & Weisbuch, 2010; Barrett, Mesquita, & Gendron, 2011), we expect that no single cue will possess substantial trust-related predictive power on its own. As Keltner and colleagues have shown, nonverbal signals of complex social states like embarrassment are often composed of a set of cues (Keltner & Buswell, 1997). Indeed, sets, by their nature, can be more informative than the presence of any of their individual components because of their ability to resolve ambiguities inherent in the interpretation of single cues (cf. Hall, Coats, & Smith Lebeau; 2005). Accordingly, we expect any trust-relevant signal to be composed of sets of cues emitted in close temporal proximity that, when taken together, convey an individual's intentions.

To examine if and how individuals can assess whether a novel partner will cooperate in a trustworthy manner, we designed a two-phase strategy. The goals of the first phase were two-fold: to demonstrate that exposure to nonverbal cues increases accuracy in assessing trustworthiness and to identify a set of cues that are reliably predictive of trust-relevant behavior. The goal of the second phase was to manipulate the expression of the candidate cues with exacting precision in order to assess their impact on subsequent decisions to trust a partner.

Experiment 1

Experiment 1 constituted the exploratory phase of this project. The primary goal was to identify a set of nonverbal cues that constitute a signal related to the trustworthiness of a novel partner. To accomplish this goal, we constructed a paradigm in which individuals would interact with a previously unknown other in a "get-to-know-you" conversation (either face-to-face or over a web-based chat), following which they would play an economic game with this individual that pitted self-interest against joint-interest.

The strategy in this phase was quite straightforward. First, if information about the trustworthiness of another is carried through nonverbal cues, then accuracy in judging the cooperative intent of a partner should be greater when interactions occur face-to-face, as opposed to over a web-based chat where only semantic information is available. Second, if such nonverbal information exists, we should be able to identify a candidate cue set by linking expressions of specific cues to actual economic behavior. Identification of such a signal would, of course, represent an initial step, with final confirmation of a cue set requiring validation through experimental manipulation in Experiment 2.

Method

Participants. Eighty-six individuals (34 male, 52 female) from the undergraduate participant pool at Northeastern University agreed to take part in the experiment. These individuals were assigned randomly to 43 dyads, with the only requirement for assignment being unfamiliarity with the assigned partner.

Procedure. As noted, the procedures for the experiment are divisible into two distinct phases: the initial interaction followed by the economic decisions.

Initial interaction. Dyads were randomly assigned to one of two conditions: face-to-face interaction (FTF) or web-based-chat (WBC). In the FTF condition the participants were brought into a single lab and seated at a table. In the WBC condition the two participants were instead brought individually into two separate rooms. Participants in both conditions were told that the purpose of the study was to explore how people form impressions of one another, but were not told any details about the second half of the experiment which involved an economic exchange game.

In this first part of the experiment, participants were asked to have a conversation for 5 minutes. Participants in the FTF condition spoke face-to-face whereas participants in the web-based-chat condition spoke over the Internet using AOL instant messenger (AIM). Participants using AIM were asked to refrain from using emoticons. In both conditions, participants were encouraged to speak about whatever they liked, with the exception of what tasks might be coming next. They were given several conversation probes (e.g., *What are your plans for spring/summer break? What do you like about living in Boston?)* but were told that they should not feel limited to these topics.

The reason for prohibiting any discussion of the upcoming economic game was to remove any possibility for strategic deception. Given that individuals could not talk about the game, or even know its rules or form, active deception was unlikely. Rather, it was our goal to determine if a partner's general level of cooperative intent could be discerned prior to engaging in any type of negotiation with them.

After providing the instructions, the experimenter left the participants alone to have their conversation for 5 minutes. Participants in the FTF condition were recorded digitally for the duration of their conversation from three different cameras: two that captured a head-on view of each participant and one that captured a side view of the dyad as a whole. In the WBC condition, participants' AIM dialogue was recorded as text.

Economic decisions. Following the initial interaction, participants played a single round of the Give-Some-Game (GSG) (cf. DeSteno, Bartlett, Baumann, Williams, & Dickens, 2009; van Lange & Kuhlman, 1994). Participants in the FTF condition were removed to separate rooms to complete the GSG; those in the WBC condition remained in their separate rooms to complete it. The GSG is an analog of a typical prisoner's dilemma but allows for a range of cooperative vs. self-interested behaviors. Each participant was given four tokens, each worth $1 to her if she kept it, but $2 if given to the partner. Maximal cooperation and communal gain occurs when each individual gives all four tokens to her partner, resulting in an $8 payoff for both. Maximal individual (i.e., selfish) gain occurs when one partner gives all four tokens and the other gives none, resulting in a payoff of $12 for the receiver and $0 for the giver. The GSG thus allows for incremental measures of cooperative intent with intermediate levels of giving corresponding to degrees of cooperative (i.e., trustworthy) or selfish (i.e., untrustworthy) behavior.

In addition to reporting how many tokens they would offer to their partner, participants also reported their estimates of the number of tokens they believed their partner would offer them. Once offers were complete, participants received their respective payoffs. There was no expectation that partners would see each other again; therefore, individuals had no reason to feel a pressure to reciprocate.

Coding of Nonverbal Cues. The digitally-recorded interactions from the face-to-face condition were coded by a set of independent coders, with each interaction being coded by 2 separate individuals (interrater agreement: Rho = .87). Using all three camera angles and Noldus Observer XT software, coders marked start/stop times for each of 12 types of cues throughout each interaction. The cues to be coded were selected based upon their frequencies of appearance; a cue had to be expressed by at least 5 participants in the data set (although in practice, most were expressed by many more individuals) to be coded. The result provided a time-synchronized stream of nonverbal cues that were emitted by each participant in a dyad. The resulting 12 nonverbal cues in the coding scheme were the following: smile, laugh, lean forward, lean back, arms crossed, arms open, face touch, hand touch, body touch, head shake, head nod, look away.[3]

Results

A mixed-model ANOVA treating dyad as a random factor to control for nested dependencies among participants confirmed that, in accord with our primary hypothesis, accuracy in predicting trustworthy behavior was greater when individuals had access to the nonverbal cues of their partners, $F(1, 41) = 3.99$, $p = .05$. As depicted in Figure 1, smaller average prediction errors (i.e., smaller average absolute values of the discrepancies between predicted and received tokens) occurred in the FTF condition than in the WBC condition. Of import, general levels of giving were equivalent in both groups (see Figure 1), demonstrating that access to nonverbal cues enhanced accuracy in assessing subsequent trustworthy behavior but did not influence the actual occurrence of such behavior itself.

In order to identify a candidate signal that was predictive of trustworthy or untrustworthy behavior, we sought to construct candidate sets of cues based on both existing knowledge relating cues to affiliative or avoidant behavior (Ambady & Weisbuch, 2010; Knapp & Hall, 2010) and examination of zero-order correlation matrices for associations linking frequencies of 12 different observed cues with actual economic behaviors.[4] We examined the ability of combinations of cues to predict cooperative behavior using multilevel models that allowed for the control of dyadic dependencies within the data. The model used to assess the signal value of cue sets with respect to self- and partner-expressed cues, respectively, took the following form:

$$\widehat{Y_{ij}} = \beta_{0j} + \beta_{1j}X + r_{ij}$$

$$\beta_{oj} = \varphi_{00} + \mu_{0j}$$

$$\beta_{1j} = \varphi_{10}$$

Here, Y_{ij} refers to the number of tokens offered for exchange by participants (i) nested in dyads (j); X refers to the mean of the frequencies of the cues (either emitted by oneself or emitted by one's partner); r_{ij} refers to participant level error; φ_{00} and φ_{10}refer to population-value estimates for the respective intercepts and slopes linking frequencies of the cues to the tokens given; and μ_{0j} refers to dyad-level variability in intercept values of the dependent variable. The final candidate set for a trust-relevant signal, determined based on prediction power for cooperative behavior and established knowledge relating cues to affiliation- and avoidant-relevant intentions, consisted of four cues: hand touching, face touching, crossing arms, and leaning away.

As expected, none of these cues when examined in isolation offered significant predictive ability. However, when taken together in a unit-weighted manner (i.e., the mean value across the set of four cues) the resulting signal was predictive of trust-relevant behavior. The more frequently an individual expressed these cues, the less trustworthy he or she behaved (i.e., the fewer tokens he or she offered for exchange), φ_{10}= -0.15, *p* = .03. Similarly, the more frequently an individual's partner expressed these cues, the fewer tokens he or she decided to offer the partner, φ_{10}= -0.13, *p* = .03.

Discussion

The findings of Experiment 1 demonstrate that exposure to nonverbal information increased accuracy in assessing the trustworthiness of another by approximately 30%, thereby supporting the existence of a trust-relevant signal. In addition, we were able to identify a specific candidate cue set that was directly associated with trust-related economic behavior. It is important to note, however, that although the identified cue set received empirical validation in this sample, the nature of the study was clearly exploratory. As such, the findings could have stemmed from random sample variations or from spurious correlations between specific cues. For example, given that most people unconsciously generate a multitude of potential cues, certain cues might co-vary with others (e.g., nodding a head while leaning away), making spurious correlations between some subsets of cues nearly inevitable. In order to validate the cue set, we therefore needed to achieve exacting control over all potentially relevant cues so that we could manipulate and isolate them experimentally. This strategy would serve as a truly confirmatory endeavor and allow for a clean test of causality for the cue set.

Experiment 2

Experiment 2 constituted the confirmatory phase of this project. Specifically, its goals were to allow for experimental validation of the causal efficacy of the target cue set identified in Experiment 1. Accomplishing this goal required the ability to manipulate target cues orthogonally to nontarget ones with high precision. However, a fundamental challenge inherent in this research design is that individuals regularly emit cues outside of their own awareness, which makes it difficult even for trained confederates to express individual cues in a reliable and orthogonal fashion. Our strategy for meeting this challenge was to employ a social robotics platform that allowed the specific cues emitted by one member of the dyad (the robot) to be controlled to a degree not possible with humans.

The resulting procedures for Experiment 2 thus closely mirrored those of Experiment 1, with the principal exception being that one of the members of each dyad was replaced with the robot Nexi (see Figure 2). The primary manipulation centered on whether Nexi expressed the target cue set identified in Experiment 1. The causal power of the cue set would be confirmed to the extent that Nexi's expression of the target cues resulted in reduced perceptions of the robot's trustworthiness and subsequent economic behavior toward it.

Method

Participants. Sixty-four individuals from the greater Boston community agreed to take part in the experiment (22 male, 42 female; mean age = 21, sd = 2.06 years). These individuals were randomly assigned to one of two conditions: receipt of target cues vs. control.

Procedure. The procedure for the second experiment was based on that of the first, but with several noteworthy exceptions. The primary difference was that we employed a Wizard-of-Oz paradigm, so called because Nexi was controlled by two operators in a separate room. Participants completed the experiment individually as opposed to in dyads (i.e., Nexi was the partner for each participant). They were brought into a lab where the humanoid robot was already positioned. Participants were seated across from Nexi with a small, low table separating them. Nexi welcomed each participant with a wave, saying "Hello. It's nice to meet you." Participants were then told that for the first part of the experiment they would have a conversation with Nexi for 10 minutes. As before, they were not given any specific details about the second half of the experiment which involved playing the GSG. The duration of the conversation was extended from the previous study's 5 minutes to allow participants time to adjust to interacting with a robot. At the start of the conversational segment, Nexi provided a few details about herself (e.g., where she was built) to allow participants to become comfortable with conversing with a robot.

Participants were again given a set of conversation probes, but in this experiment they were asked to stick to the topics provided during their interaction (e.g., *What do you like about living in Boston?*). Participants were not told that Nexi was being tele-operated by experimenters in an adjacent room. The experimenter left the participant in the room with Nexi for the duration of their interaction. Three different cameras were again used to record all interactions.

Thirty-one of the 64 participants were assigned to the target cue condition; the remaining 33 were assigned to the control condition. In the control condition, Nexi made several conversational gestures throughout the interaction, but did not engage in any of the target cues that were found to be predictive of untrustworthy behavior in Experiment 1. In the experimental condition, some of the conversational gestures from the control condition were replaced with the target cues (hand touch, face touch, lean back, and arms crossed), ensuring an average of 1-3 occurrences of each target cue (i.e., frequencies similar to occurrences within individuals in Experiment 1). The robot's expression of all cues was based on prototypical analogs of human motions. That is, the gestures were created by an animator who distilled them from examples of human motions.[5] Care was taken to keep the robot's overall amount of movement consistent across conditions.

Prior to beginning the experiment, researchers prepared scripted responses to each of the conversation probes. Nexi was tele-operated by two experimenters in an adjacent room. A single, female experimenter served as the voice of Nexi and interacted with all participants. Her speech and head movements were tracked and mapped to the robot's head and mouth movements in real-time. Cameras and microphones embedded in the robot allowed the experimenters to see and hear participants conversing with the robot, and face-tracking software was utilized to keep Nexi's gaze centered on participants' faces. A second experimenter operated Nexi's torso and arms throughout the conversation using a graphical user interface (GUI). In this way, the experimenter speaking with the participant could remain blind to the experimental condition (i.e., target cues versus control). Detailed specifications for the technology underlying Nexi's animation and control, as well as video examples of its interaction with participants can be found in the SOM-U.

Following the "get-to-know-you" period, participants were moved to a separate room where they played the GSG assuming Nexi was their partner. Participants also completed several questionnaires that probed their views of and familiarity with Nexi, including individual items assessing how much they trusted Nexi and how much they liked the robot using 7-point scales.

Results

The causal efficacy of the signal in question would be confirmed to the extent that Nexi's expression of cues would cause participants to judge it to be less trustworthy and correspondingly lead them to expect and offer fewer tokens within the economic game. We used the structural equation model depicted in Figure 3 to examine the validity of this hypothesis. Confirming the cues' signal value, Nexi's presentation of the target cues produced a significant decrement in her perceived trustworthiness. Of import, this decrement directly predicted lower expectations for the number of tokens Nexi would give and for the number of tokens participants actually offered to exchange with it. Of note, presentation of these cues had no effect on economic decisions other than via its impact on perceived trustworthiness. Trimming the direct paths linking condition to the number of tokens expected to be received and the number of tokens given did not diminish the model's goodness of fit, $\chi^2(2) = 3.76$, $p = .15$.[6]

Finally, the effects of the cues appear quite narrowly focused on trust, as their presence or absence did not influence the degree to which participants liked Nexi ($t < 1$). This finding suggests that the presence of the cues did not produce a general "anti-halo" effect such that evaluations for the robot on any social dimension became more negative. In many ways, this finding mirrors a familiar experience for many people in that most can point to individuals who they like but with whom they would not trust their money.

Discussion

These findings are noteworthy for two primary reasons. First, they provide a stringent confirmatory test for the nonverbal signal identified in Experiment 1. The robotic system allowed us to gain precise control over the cues in order to manipulate them in an experimental manner (i.e., as orthogonal from any other cues) and, thereby, provide a test of their causal impact. As predicted based on the correlational data from Experiment 1, presence of the cue signal caused individuals to perceive Nexi to be less trustworthy, which directly impacted their economic behavior toward it.

Second, these findings also offer the first evidence that the human mind will respond to trust-relevant signals emitted by humanoid robots in the same manner as to similar signals emitted by humans. It remains to be explored whether the impact of these cues stems from the ascription of moral intentions to Nexi's "mind" (cf. Gray, Young, & Waytz, in press; Waytz, Epley, & Cacioppo, 2010; Waytz, Gray, Epley, & Wegner, 2010), or simply from utilization of the cues in a nonconscious manner as predictors of Nexi's subsequent moral behavior. Irrespective of mechanism, however, these findings clearly indicate the readiness of the mind to respond in the expected manner to human-like biological motion emitted by robotic entities.

General Discussion

Taken together, these findings are among the first to identify a human capacity to assess whether another unfamiliar individual is likely to behave cooperatively in a given situation. As such, they offer empirical support for a phenomenon that has long been theorized to allow for the existence of cooperation, especially in one-shot dilemmas. Yet, it is also important to note that the cues we have identified are almost surely not the only ones that affect judgments of trustworthiness through their expression or suppression. Unlike the circumstances examined here, in which cooperation was fairly common, the default expectation in many contexts may be that cheating will occur. In such situations, a different subset of cues, such as those typically associated with motivations for affiliation (e.g., leaning forward, affirmative head nods), might hold greater predictive power to detect fair and cooperative, as opposed to selfish and opportunistic, tendencies. Indeed, the informational value of any set of cues is likely to depend upon context-based differences in expectancies for the intentions of partners.

As noted, these findings also offer initial evidence that the human mind will use nonverbal cues to predict the trustworthiness of humanoid robots, thereby opening many avenues for increasing the capacity of robots to function as interaction partners capable of building trust and social bonds with humans through either the presence or absence of specific gestures. In so doing, they support the view that robotics technology has now reached a level where its mirroring of human social cues, though imperfect, is nonetheless sufficient to embody the basic components necessary to engage the social mind's interpretive machinery.

We readily acknowledge the view held by some that robotics might not constitute a valid method to study human behavior, as robots clearly do not look or move exactly like their human counterparts. Individuals who harbor these concerns worry that such imperfections might bias the human mind's responses to nonverbal cues. Although this is certainly a valid concern, it is one that, to our minds, needs to be answered empirically as opposed to being based on subjective impressions of the "humanness" of any robot. Indeed, within the paradigm of "computers as social actors" a wide range of technological embodiments capable of expressing social cues (from desktop computers that interact via text, to animated avatars, to physical robots) has been shown to evoke natural human social responses and social judgments (citations?).

Had Experiment 2 failed to confirm the findings of Experiment 1 in the present case, any number of reasons could have been posited. For example, the findings from the human-human interactions could have been incorrect or the technology of the robot might not have been capable of adequately mirroring human movement. However, when the findings of the experiments are viewed as a whole in that manipulation of the cues by the robot confirmed the cues' predicted impact based on human-human interactions, logic clearly dictates that the most parsimonious explanation is the proffered one. To believe otherwise requires one to embrace the view that when a robot expressed the trust-related cues identified in Experiment 1 to an entirely different sample, a biasing agent (e.g., timing of movements) inserted itself that not only produced behavioral effects that matched those from the human-human interactions, but also did so via the predicted mediator. This would constitute a procrustean explanation indeed.

To our mind, the use of interdisciplinary techniques can often be accompanied by increased ambiguities in the interpretation of any single experiment. The simple act of combining different technologies and paradigms can similarly combine the methodological shortcomings of each, thereby increasing the number of ways failures to replicate can occur. However, when findings converge across methodologies, confidence in their robustness is greatly increased, as the number of ways replication can erroneously emerge becomes vanishingly small.

Notes

[1]Delton and colleagues have presented a model simulation showing that cooperation in one-off interactions can occur regularly if expectations for future interaction are non-zero. However, their model also suggests presence of an adaptive advantage in deciding to cooperate stemming from an ability to gain insight into the trustworthy intent of a partner.

[2]One exception has been evidence suggesting that individual differences in facial width are associated with greater selfish behavior (Stirrat & Perret, 2010). Such a marker, however, is static, and thus unable to account for differences in situation-by-situation interactions.

[3]Table S1 in SOM-R provides descriptive information regarding cue frequencies.

[4]See SOM-R for more detailed information on cue selection.

[5]Prototypical examples of target cues were derived from the video recordings of Experiment 1.

[6]Note that the initial model is saturated, and thus fits perfectly. Therefore, the nonsignificant Chi-square test for the reduced model indicates not only a lack of substantive change in fit for the reduced model, but also an acceptable fit for the reduced model overall.

References

Ambady, N., & Weisbuch, M. (2010). Nonverbal behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th Edition, pp. 464-497). Hoboken, NJ: John Wiley & Sons.

Axelrod, R. (1984). *The Evolution of Cooperation.* New York, NY: Basic Books.

Barrett, L., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, 20, 286-290.

Delton, K. W., Kraznow, M. X., Cosmides, L. & Tooby, J. (2011). The evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences, 108*, 13335-13340.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions* . New York, NY: W. W. Norton & Company.

Gray, K., Young, L., & Waytz, A. (in press). Mind perception is the essence of morality. *Psychological Inquiry.*

Hall, J. A., Coats, E. J., & Smith-Lebeau, L. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin, 131, 898-924.*

Keltner, D., & Buswell, B. N. (1997). Embarrassment: Its distinct form and appeasement functions. *Psychological Bulletin, 122,* 250-270.

Knapp, M. L., & Hall, J. A. (2010). *Nonverbal Communication in Human Interaction* (7th Edition). Belmont, CA: Wadsworth.

Stirrat, M., & Perrett, D. L. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21,* 349-354.

Todorov, A. (2008). Evaluating faces on trustworthiness. *Annals of the New York Academy of Sciences, 1124,* 208-224.

Todorov, A., Baron, S., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social, Cognitive, and Affective Neuroscience, 3,* 119-127.

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social cognition unbound: Psychological insights into anthropomorphism and dehumanization. *Current Directions in Psychological Science, 19,* 58-62.

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences, 14,* 383-388.

Figure Captions

Figure 1. Mean Absolute-Value Prediction Errors in Expected Number of Tokens from Partner (Panel A) and Mean Number of Actual Tokens Given to Partner (Panel B) for Participants in Experiment 1. Error bars signify ± 1 standard error.

Figure 2. Nexi the humanoid robot tele-operated in Study 2.

Figure 3. Path model depicting economic decisions as a function of target cue exposure and perceived untrustworthiness. Standardized parameters are presented. Asterisks indicate P's ≤ .05. Condition is dummy coded (0 = absence of target cues, 1 = presence of target cues).